

Title	音声信号と調音状態の一对多の関係の分析及びその応用に関する検討
Author(s)	錦戸, 信和
Citation	
Issue Date	2011-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/9605
Rights	
Description	Supervisor: 党建武, 情報科学研究科, 博士

音声信号と調音状態の一对多の関係の分析 及びその応用に関する検討

錦戸 信和

北陸先端科学技術大学院大学

2011年3月

論文の内容の要旨

音声生成過程において、声道を形成する各調音器官の様々な位置や形状、すなわち調音状態から同一の音素カテゴリーに含まれる音響特徴量を持つ音声信号を生成することが可能である。この音声信号と調音状態の一对多の関係は、音声信号から調音状態を逆推定する場合に不良設定問題、つまり入力音声に対して逆推定の解となる調音状態が無数に存在する一对多の問題を生じさせる。従来の研究は、形態学的制約条件、動的制約条件、生理学的制約条件を導入し解の多様性を減少させることで一对多の問題の解決を試みているが、完全な解決には至っていない。ここで、腹話術や調音の補償動作の観測実験から、音声信号と一对多の関係にある調音状態には、自然な発話を行う際に観測され得る“自然調音状態”と、生理学的に可能な状態であるが自然な発話を行う際に観測され得ない“不自然調音状態”が含まれると考えられる。この不自然調音状態は、自然調音状態と同様の音響特徴量を持つ音声信号の生成が可能であるが通常の発話の際には観測され得ないため、逆推定の解の調音状態に多様性をもたらす要因となる。しかしながら、多くの不自然調音状態は従来の逆推定の制約条件を満たすため、逆推定の推定候補に含まれる不自然調音状態を取り除くことが、一对多の問題を解決するために必要と考えられる。

本研究では、音声信号から調音状態を逆推定する際に推定候補に含まれる不自然調音状態を除去することを目的とする。そのために、発話の計算モデルを用いて人間が発話可能な多数の多様な調音状態を生成し、生成した調音状態の分析に基づく分布構造の可視化により、音声信号と一对多の関係にある人間が発話可能な調音状態の分布の全体像を明らかにすることを試みた。さらに、調音状態の分布構造に基づき自然調音状態と不自然調音状態を識別する手法を提案し、提案手法を音声信号から調音状態を逆推定するシステムに適用することにより、逆推定の推定候補に含まれる不自然調音状態の除去を試みた。

上記の結果から、本研究では以下の成果が得られた。

1. 音声信号と一对多の関係にある人間が発話可能な調音状態の分布の全体像を、自然調音状態と不自然調音状態を含む非線形空間上のクラスタ構造として示した。
2. 調音状態のクラスタ構造に基づき自然調音状態と不自然調音状態を識別する手法を提案し、自然調音状態に対して97%、不自然調音状態に対して99%の精度で識別可能なことを示した。
3. 提案した識別手法を音声信号から調音状態の逆推定に適用した新たな逆推定システムを構築した。
4. 構築したシステムによる逆推定の結果、推定候補に含まれる不自然調音状態を9割除去可能なことを示した。

本研究で得られた上記の成果は、音声信号から調音状態の逆推定に大きく寄与するだけでなく、聴覚障害者や語学の学習者のための理想的な発話訓練システムの実現に貢献できると考えられる。また、本研究で得られた知見は、人間の音声生成機構の解明や、音声合成の研究に大きく寄与できると考えられる。

キーワード: 一对多の問題, 逆推定, 不自然調音状態, 調音モデル, 母音生成